PTO/PCT Rec'd 10 JAN 2001
09/743533

Modified Proteins

Technical Field

The present invention is directed generally to producing modified proteins and particularly to producing modified glutenin or seed-storage proteins.

Background Art

Wheat storage proteins are classified on the basis of their solubility into two classes. The gliadins are readily soluble in aqueous alcohols and are monomeric proteins with only intramolecular disulphide bonds. The glutenins are present in high molecular weight polymers, stabilised by intermolecular disulfide bonds and are not soluble in aqueous alcohols without reducing agent (Kasarda 1989). These proteins are present in high amount in the endosperm and are considered to act as a store of nitrogen, carbon and sulphur for seed germination.

Glutenins form a continuous proteinaceous network called gluten. The unique physico-chemical properties of gluten determine the ability of wheat dough to be processed into baked goods (bread, biscuits, cakes), pasta noodles and other food products. It is understood that the glutenins, which form crosslinks with each other through disulfide bonds, are the most important molecules producing the viscoelastic properties of wheat flour dough (MacRitchie 1992). The unique position of wheat in bread making is due to the ability of the dough to retain gas on expansion. The gluten accounts for about 10% of the dough, and consists mainly of proteins (70-80%) together with starch and lipids. Starch could be granular and damaged starch. The lipid reserves of wheat are non-polar, structural and endosperm lipids (Gan et al., 1995). Structural lipids are also called polar lipids. The endosperm lipids are divided into non-starch lipids and starch lysophospholipids. The structure and properties of gluten are determined by molecular interactions and it is important that these be understood if the functional properties of gluten are to be manipulated.

A dough results from a large variety of interactions between flour constituents facilitated by water. Starch takes up about 46% of the water and damaged starch contributes significantly to the water absorption. It has been shown that during hydration, proteins exude visible strands or fibrils. Specific proteins of flour are bound to flour lipids (polar) upon addition of water (Morrison 1989).

Dough development is visualised as a re-orientation of glutenin polymers to form a membrane network with viscoelasticity and gas retaining properties. Covalent (disulfide) and noncovalent (hydrogen, hydrophobic and ionic) bonds are involved in formation of a fully developed dough. Interactions are further modified during fermentation, baking and even after baking. The disulphide bonds of flour proteins play a key role in the interactions in dough. The bonds form relatively strong crosslinks within and between polypeptide chains and also stabilise other less energetic bonds. Disulphide bonds provide the required stability for the protein matrix until the loaf structure is set by the gelatinisation of the starch and the thermal denaturation of the proteins during baking. Hydrogen bonds are considerably weaker than covalent bonds, but contribute significantly to the structure of dough. A unique feature is the ability to interchange with other hydrogen bonds, which facilitate reorientation of protein chains and allow for stress relaxation. Hydrophobic bonds result from nonpolar groups of flour constituents. Because these bonds are reversible, they can readily accommodate viscous flow and thereby facilitate mechanical dough development. Ionic bonds play relatively small part in dough structure formation but some specific components have an ionisable part or parts. Therefore ionic bond interactions could be important for the rheological properties (for a review, see Bushuk 1998).

A major limitation to evaluating the contributions of various groups of proteins, and of specific structural features of these molecules, to dough functionality has been the lack of appropriate systems that allow specific proteins to be incorporated and tested within the dough. The situation has recently changed, however, due to two advances. The first is the development of small scale testing equipment (Mixograph, Extensograph) with appropriate procedures for incorporating exogenous proteins, including polymeric glutenins into the dough (Bekes et al., 1994). Advantages of this system are the small amount of proteins required for test and the ability to rapidly test multiple samples produced by, for example, protein engineering. The second recent advance is the development of a reliable transformation system for wheat (Weeks et al., 1993, Witrzens et al., 1998), which allows the modification of storage protein composition by the expression of new proteins with, for instance, designed characteristics.

PCT/AU99/00563
Received 1 June 2000

ART 34 AMDT

To alter protein-protein, protein-lipid and protein-starch interactions within the gluten matrix, the present inventors have developed a system which enables the incorporation of new surface active molecules or parts of molecules into the gluten matrix.

Disclosure of Invention

In a first aspect, the present invention consists in a method for producing a modified glutenin or seed-storage protein, the method comprising adding to the protein an exogenous amino acid domain which confers to the modified protein the ability to bind a ligand or other macromolecule, wherein the modified protein has an ability to incorporate into gluten.

In one preferred embodiment, the modified glutenin or seed-storage protein further contains one or more exogenous amino acid residues added to its amino acid sequence. More preferably, the one or more amino acid residues are one or more cysteine residues. Preferably, the one or more cysteine residues are incorporated at one or both ends of the amino acid sequence of the protein. The addition of the one or more cysteines allows the modified proteins to be more easily incorporated into gluten in use. The further modifications to the glutenin or seed-storage proteins produced according to the present invention allow or can assist in the incorporation of that protein into the gluten network for food or industrial use.

The present inventors have found that incorporating exogenous amino acid sequences (domains) from proteins other than glutenins or seed storage proteins into glutenin or seed-storage proteins modifies the general properties of gluten, particularly when the modified proteins are used in a range of food applications.

Figure 16 provides a schematic of the scheme for identifying transgenes in transformed wheat plants by polymerase chain reaction (PCR). A primer pair straddling the interface between the gene and its promoter (from the gene for the high molecular weight glutenin Bx17 for example) ensures that no false positives are detected arising from the high homology between C hordein and gliadin genes.

The binding domain can be any domain that will bind ligands that may be useful in food preparation or in food compositions. In a preferred form, the binding domain is a ligand capable of binding lipids or starches. The

present inventors have found that the lipid-binding domain of the barley oleosin gene, the lipid-binding regions of the wheat CM16 protein, and the starch-binding domain of the glucoamylase from *Aspergillus niger* are particularly suitable for the present invention. It will be appreciated, however, that other natural or modified domains would also be suitable for the present invention.

One glutenin or seed-storage protein that has been modified by the present inventors is the C hordein gene from barley. It will be appreciated, however, that other glutenin or seed-storage proteins may also be modified according to the present invention. In wheat, such glutenin or seed proteins include low molecular weight glutenins, high molecular weight glutenins, gliadins, puroindolines or grain softness proteins (also known as friabilins), or Chloroform/Methanol-soluble proteins. Homologues of these proteins exist in other cereals such as diploid, tetraploid and hexaploid wheats, rye, triticale, barley, oats, rice, sorghum, millet and maize and the genes encoding these proteins may also be modified according to the present invention.

In a second aspect, the present invention consists in a modified glutenin or seed-storage protein having a domain inserted therein which confers to the protein the ability to incorporate into gluten or bind a ligand or other macromolecule.

In one preferred embodiment, the modified glutenin or seed-storage protein is produced by the method according to the first aspect of the present invention.

In another preferred embodiment, the modified glutenin or seed-storage protein is ANG/SBD/Cys7Cys236, ANG/OHBD/Cys7Cys236 or ANG/CM16/Cys7Cys236.

In a third aspect, the present invention consists in an isolated nucleic acid molecule encoding a modified glutenin or seed-storage protein according to the second aspect of the present invention.

In a fourth aspect, the present invention consists in an isolated nucleic acid molecule according to the third aspect of the present invention incorporated into a cell such that on expression of the nucleic acid molecule, the cell produces the modified glutenin or seed-storage protein.

The cell may be a recombinant bacterial cell, for example, which is capable of producing the modified glutenin or seed-storage protein. Preferably the bacterial cell is *Escherichia coli*. Alternatively, the cell may be

a yeast such as *Pichia* sp. or *Saccharomyces cerevisiae*, an insect cell using an expression system such as the baculovirus expression system, or a mammalian cell. Alternatively, the cell may be a plant cell of a recombinant plant which is capable of producing the modified glutenin or seed-storage protein in the plant's seeds. Preferably the plant cell is a recombinant wheat cell.

In fifth aspect, the present invention consists in the use of a modified glutenin or seed-storage protein according to the second aspect of the present invention in the preparation of a food product.

Examples of food products include leavened or unleavened breads, pasta, noodles, breakfast cereals, snack foods, cakes, pastries or other foods containing flour-based sauces or ingredients.

The modified glutenin or seed-storage proteins according to the present invention, in use, are capable of modifying the structure of doughs and other materials containing gluten in ways which add value and utility to the resultant product. The modified glutenin or seed-storage proteins are suitable for use in the food industry as modifiers of food properties.

The present inventors have shown that modified proteins according to the present invention can be produced in bacterial fermentation and that large scale production of the proteins for commercial use is possible.

In sixth aspect, the present invention consists in the use of a modified glutenin or seed-storage protein according to the second aspect of the present invention in the preparation of a non-food product.

Examples of non-food products include, but not limited to, films, coatings, adhesives, building materials or packaging materials. It will be appreciated that the modified proteins according to the present invention would have the same non-food uses as for normal glutenin or seed-storage proteins.

In a seventh aspect, the present invention consists in the use of a grain or part of a grain containing a modified glutenin or seed-storage protein according to the second aspect of the present invention in the preparation of a food product.

It will be appreciated that the modified glutenin or seed-storage proteins according to the present invention may be contained in, or produced by, a transgenic plant produced by the fourth aspect of the present invention. Thus, the grain or other plant products produced by these plants may be used

in the crude form of flour, semolina, bran, pollard, germ fraction, or the like. Such plants may produce a mixture of normal and modified glutenin or seed-storage proteins which can also be used. Mixtures of modified glutenin or seed-storage proteins and other plant-based materials can also be prepared to form improved gluten.

Throughout this specification, unless the context requires otherwise, the word "comprise", or variations such as "comprises" or "comprising", will be understood to imply the inclusion of a stated element, integer or step, or group of elements, integers or steps, but not the exclusion of any other element, integer or step, or group of elements, integers or steps.

In order that the present invention may be more clearly understood, preferred forms will be described with reference to the following examples and drawings.

<u>Brief Description of Drawings</u>

Figure 1: Sequence of the pJANGΔCys7Cys236 vector (SEQ ID NO: 14 and SEQ ID NO: 15). Restriction sites and enzymes of the MCS of pJKKm and pGEM-T plasmids are written in bold. Cloning sites for gene insertion into ANG are underlined and written in bold. Cleavage sites for sub-cloning the gene into pET 11d expression vector are underlined.

Figure 2: Nucleotide (SEQ ID NO: 16) and amino acid sequence (SEQ ID NO: 17) of Oleosin Hydrophobic Binding Domain (OHBD). Arrowheads are indicating the direction of primer extension. Smaller letters are indicating extra nucleotides for cloning and amino acids from ANG molecule. Amino acid sequence of the central anti-parallel domain is designated by two arrows.

Figure 3: Nucleotide (SEQ ID NO: 18) and amino acid sequence (SEQ ID NO: 19) of ANGΔCys7Cys236 (molecular mass 18.5 kDa) Cysteine residues are marked with bold letters. Extra two amino acids and six nucleotides are written with smaller letters.

Figure 4: Nucleotide (SEQ ID NO: 20) and amino acid sequence (SEQ ID NO: 21) of the starch binding domain of 1,4-α-D-glucan glucohydrolase from *Aspergillus niger* (molecular mass 11.9 kDa). Small letters are indicating extra nucleotides for cloning and amino acids from ANG molecule. Restriction sites are in bold and underlined.

Figure 5: Nucleotide (SEQ ID NO: 22) and amino acid sequences of CM16 (SEQ ID NO: 23) and CM17 (SEQ ID NO: 24) (molecular mass is 13.4

kDa). Small letters are indicating extra nucleotides for cloning and amino acids from ANG molecule. Restriction sites are in bold and underlined. First amino acid sequence under the nucleotide sequence represents the CM16 protein, while only differences are shown in CM17 protein.

Figure 6: Nucleotide (SEQ ID NO: 25) and amino acid sequence (SEQ ID NO: 26) of Puroindoline A (molecular mass is 14.3 kDa). Small letters are indicating extra nucleotides for cloning and amino acids from ANG molecule. Restriction sites are in bold and underlined.

Figure 7: Expression of the ANG/OHBD/Cys7Cys236 gene. SDS-PAGE analysis followed by Coomassie blue staining. Lane A contains standard molecular weight markers. Lane B shows a control in which the host *E. coli* cell contains the pET-11d vector. Lane C shows the expression of the recombinant ANG/OHBD/Cys7Cys236 protein (migrating slightly further than the 30 kDa marker) from pET-11d containing the ANG/OHBD/Cys7Cys236 gene.

Figure 8: Expression of the ANG/OHBD/Cys7Cys236 gene. SDS-PAGE analysis followed by Western blot using antibodies to C-hordein. Lane A contained standard molecular mass markers. Lane B contained extract of an *E. coli* cell expressing a modified C-hordein gene containing a single cysteine in the N-terminal region. Lane C contained extract from *E. coli* containing the ANG/OHBD/Cys7Cys236 gene in the pET-11d expression plasmid.

Figure 9: SDS-PAGE analysis of the purified ANG/CM16/Cys7Cys236 gene product by Coomassie blue staining. Lane A contains standard protein molecular weight markers. Lanes B to E show ethanol-soluble extracts of the crude *E. coli* lysates. Lane B contained extract from cells containing the control plasmid, pET-11d. Lanes C, D and E contained ethanol soluble extracts of cells harbouring the pET-11d vector containing the ANG/CM16/Cys7Cys236 gene, prepared from cells 2, 4 and 6 hours, respectively, after induction of protein synthesis using IPTG.

Figure 10: Expression of the ANG/PIN-A/Cys7Cys236 gene. SDS-PAGE gel stained with Coomassie blue. Lane A contains standard protein molecular weight markers. Lane B contained extract of cells harbouring the control pET-11d plasmid, lanes C and D contained extracts of cells harbouring the pET-11d vector containing the ANG/PIN-A/Cys7Cys236 gene. Lane C contains extract 2 hours after IPTG induction, lane D contains extract prepared 6 hours following IPTG induction.

Figure 11: Expression of the ANG/PIN-A/Cys7Cys236 gene. SDS-PAGE analysis followed by Western blotting with anti-puroindoline A antibodies. Lanes A and B contained extract of cells harbouring pET-11d containing the ANG/PIN-A/Cys7Cys236 gene, 2 hours and 6 hours after induction respectively. Lane C contains a Western blot of cells harbouring a plasmid containing the puroindoline A gene alone (not inserted into the ANGΔCys7Cys236 gene). Lane D contained extract of cells harbouring the control pET-11d vector.

Figure 12: Expression of the ANG/PIN-A/Cys7Cys236 gene. SDS-PAGE analysis followed by Western blotting with anti-hordein antibodies. Lane A contains molecular weight markers. Lane B contained extract of cells harbouring the control pET-11d vector. Lane C contained extract of cells harbouring a plasmid which contains the puroindoline A gene alone. Lane D contained extract of cells harbouring pET-11d vector containing the ANG/PIN-A/Cys7Cys236 gene.

Figure 13: Plasmids for the transformation of wheat with constructs containing novel protein genes. The vectors are based on the plasmid pTLZHMWcas which has been modified by insertion of genes into the KpnI and BamH1 site region. The vectors are: pTLZ-ANGCys7Cys13, pTLZ-ANGCys236 and pTLZ_ANG/OHBD/Cys7Cys236.

Figure 14: Expression and β-cyclodextrin affinity purification of ANG/SBD/Cys7Cys236. Native- and SDS-PAGE gels of column fractions from a 5 ml β-cyclodextrin-sepharose affinity column loaded with extracts from IPTG-induced cultures of strain AD494(DE3) bearing either pET11d control plasmid or the pET-SBD plasmid for the expression of ANG/SBD/Cys7Cys236.
Panel A = SDS-PAGE of fractions from AD494(DE3)/pET-SBD.
Panel B = native PAGE of fractions from AD494(DE3)/pET-SBD.
Panel C = SDS-PAGE of fractions from AD494(DE3)/pET11d.
Panel D = native PAGE of fractions from AD494(DE3)/pET11d.
Lane M = molecular mass markers with masses indicated in kDa.
Lane 1 = the unbound flow-through from column loading.
Lanes 2-5 = four 5 ml washes with column buffer (50 mM Tris-Cl pH 7.5, 100 mM NaCl, 40 mM dithiothreitol, 1 mM EDTA).
Lanes 6-10 = five 5 ml elutions with column buffer containing 15 mM β-cyclodextrin.

Figure 15: Schematic of the method of producing a modified seed storage protein containing a binding domain for a macromolecule. In this example, the C-hordein gene is altered to produce an analogue glutenin (ANG) protein containing cysteine residues in the N- and/or C-terminal domains. This protein is then further modified by replacement of part of its repetitive domain with a binding domain from another protein. The addition of cysteine residues allows for the incorporation of the protein into, for example, the gluten macropolymer in wheat flour. The addition of a specific binding domain provides for the non-covalent linkage of the gluten network to other macromolecules (e.g. starch, lipids). Such a protein may be expressed in recombinant bacteria, yeast, or other heterologous host for use as an ingredient or additive, or it may be produced in the seed of a transgenic plant, for example wheat.

Figure 16: Schematic of the scheme for identifying transgenes in transformed wheat plants by polymerase chain reaction (PCR). A primer pair straddling the interface between the gene and its promoter (from the gene for the high molecular weight glutenin Bx17) ensures that no false positives are detected arising from the high homology between C hordein and gliadin genes.

Figure 17: Agarose gel showing PCR-based identification of a transgenic wheat plant containing a gene for a modified seed storage protein with incorporated ligand-binding domain. In this example, the gene for ANG/OHBD/Cys7Cys236 was detected in the plasmid used for transforming the wheat plant (lane 1) and in a putative transformant wheat plant (lane 2), but is not detected in a non-transformed control wheat plant (lane 3).

Figure 18: Starch-binding activity of ANG/SBD/Cys7Cys236 protein assayed by native-PAGE. Protein from each of fraction 6 and 7 from the β-cyclodextrin affinity column of the urea extract of strain AD494(DE3)/pET-SBD was loaded onto native gels containing starch (panels B and D) and/or β-cyclodextrin (panels C and D) or neither starch nor β-cyclodextrin (panel A). The reduction in electrophoretic mobility in the presence of starch alone (panel B) is indicative of the specific starch-binding activity of each protein.

Figure 19: Demonstration of the oxidative polymerisation of a modified seed-storage protein containing a macromolecule-binding domain. In this example, ANG/SBD/Cys7Cys236 was oxidised to form disulphide-linked polymeric species (indicated by arrows in the scanned image of the

SDS-PAGE gel; lanes 1, 2) in the presence or absence of added β-cyclodextrin. The polymeric species were capable of being subsequently reduced by the disulphide-specific reductant dithiothreitol (DTT; lanes 3, 4), showing that the polymerisation was due to the presence of introduced cysteine residues at the N- and C-termini of the ANG portion of the molecule.

M = molecular weight markers with masses indicated (kDa);
Lane 1 = ANG/SBD/Cys7Cys236 oxidised in the presence of β-cyclodextrin;
Lane 2 = ANG/SBD/Cys7Cys236 oxidised in the absence of β-cyclodextrin;
Lane 3 = ANG/SBD/Cys7Cys236 oxidised in the presence of β-cyclodextrin and then reduced with DTT;
Lane 4 = ANG/SBD/Cys7Cys236 oxidised in the absence of β-cyclodextrin and then reduced with DTT;
Lane 5 = reduced ANG/SBD/Cys7Cys236 in the presence of β-cyclodextrin;
Lane 6 = reduced ANG/SBD/Cys7Cys236 in the absence of β-cyclodextrin.

Modes for Carrying Out the Invention

MATERIALS AND METHODS

Bacterial strains and plasmids

Escherichia coli strain DH5α was used as cloning host strain and *E. coli* strain AD494(DE3) (Novagen) was used as the expression host in this work.

pGEM-T (Promega) was used as cloning vector of PCR products. pJKKm(-) (Kirschman and Cramer, 1988) served as cloning vehicle for fusion protein genes. New, assembled genes for expression were sub-cloned into plasmid pET-11d (Novagen).

Cloning of ANGΔCys7Cys236

Restriction endonucleases and DNA modifying enzymes were from New England Biolabs and Promega Corp. Other chemicals and reagents were of analytical reagent grade.

Oligonucleotide primers were synthesised on an Applied Biosystem 394 DNA/RNA Synthesiser using standard phosphoramidite chemistry and were deprotected by heating in ammonium hydroxide solution. Primers were lyophilised and dissolved in TE buffer.

To amplify up a 477 bp long fragment (approximately 2/3 of the whole gene) of C hordein (Accession X60037) from the genomic clone (Lambdahor1-17) of barley by polymerase chain reaction two specially designed primers were used.

oligonucleotide 1 (SEQ ID NO: 1):

```
     Met Arg                     Cys
5' GTC ATG AGG CAA CTA AAC CCT TGC AGC CAA GAG TTG CAA TC
      BspH I                                              3'
```

oligonucleotide 2 (SEQ ID NO: 2):

```
          *** Val                      Cys            Gln Gln Pro
5' GGA TCC CTA GAC CAT ACT CCA TAT GCA TGA AGC TTG TTG GGG
    BamH I                   Nde I            HindIII
GAC TGG TTG 3'
```

The reaction was performed in an FTS-4000 Thermal Sequencer (Corbett Research, Australia), with 1 cycle of 3 min at 94°C, 20 s at 55°C, 1 min at 72°C; 36 cycles of 30 s at 94°C, 1 min 30 s at 72°C. The reaction was carried out in 50 µl, containing 45 µl of Supermix (BRL Life Technology) and 1 ng of template DNA and 50 pmol of each oligonucleotides in 5 µl. The DNA was purified following the QIAquick protocol (QIAGEN) and cloned, using the pGEM-T Vector System I (Promega) as was recommended by the manufacturer. White transformant colonies selected for growth in LB medium supplemented with ampicillin (100 mg/l) were screened for insert-bearing plasmid DNA by PCR. Plasmid DNA was purified from positive clones using Jetstar miniprep columns (Genomed) and the insert was sequenced in both direction, using the Prism dye terminator cycle sequencing protocol (Perkin-Elmer). One plasmid isolate, containing the ANG gene, was designated as pGEM-ANG.

This plasmid contained two NdeI restriction sites, one within the gene and one within the pGEM-T multiple cloning site (MCS). To make it unique, the ANG gene was subcloned into pJKKm. Because pJKKm has a HindIII site in MCS, it had to be deleted before subcloning.

One ug of the plasmid was digested with 5 units of HindIII enzyme in 20 µl. Mung Bean Nuclease was used to remove the 5' overhang. Two µl of 10X $ZnSO_4$ solution and 0.02 units of enzyme was added to the reaction mix and incubated at 30°C. Reaction mix was extracted once with phenol/chloroform after one hour incubation. DNA was recovered with ethanol precipitation and resuspended in 50 µl. An aliquot of 10 µl of the DNA solution was used for ligation. Ligation was carried out in 15 µl solution at 4°C overnight using T4 DNA ligase. The ligation mixture was used to transform competent cells of *E. coli* strain DH5α by electroporation.

Colonies were grown in LB medium supplemented with kanamycin (50 mg/l). Plasmid was purified from three colonies as mentioned above. DNA samples were tested by performing HindIII enzyme digestions. One clone not containing a HindIII site was chosen and designated as pJKK-H.

pJKK-H was cut with SphI and BamHI to subclone pGEM-ANG SphI-BamHI fragment into the plasmid. Both digested DNA were purified on QIAquick columns as described earlier, and ligated in 10 μl solution overnight at 14^{0}C using T4 DNA ligase at a molar ratio of 50:1. Transformed *E. coli* competent cells was spread onto LB medium with kanamycin. Colonies were tested for insert-bearing plasmid DNA by PCR and 3 positive clones were sequenced as above. One clone was designated as pJANGΔCys7Cys236 and used for further cloning work to assemble genes for fusion proteins. Nucleotide and amino acid sequences of this cloning vehicle are shown in Figure 1.

Design and cloning of oleosin hydrophobic binding domain

The sequence of oleosin hydrophobic binding domain (OHBD) was designed to include the consensus sequence of three (maize, rice and barley) oleosin proteins. The sequence is almost identical to the sequence published for barley oleosin isoform-2 (Aalen 1995, Accession Number X82678). Four primers were designed to encode the protein (Figure 2). The gene was constructed by a modification of the technique of overlap extension, where the two long partially overlapping oligonucleotides were further extended and amplified by shorter external primers, encoding restriction enzyme cleavage sites to clone into pJANGΔCys7Cys236. PCR amplification was performed as above, using the same cycle program and Supermix solution. Concentration of the long oligonucleotides were 0.1 nM, while 2 μM of the short primers in the PCR reaction mix. The PCR product was purified on QIAquick column and cloned into pGEM-T plasmid using Promega kit, as described above. Three positive clones were used for plasmid preparation and sequencing to confirm the nucleotide sequence. One clone was used for further work and called pGEM-OHBD.

Cloning of starch binding domain

The DNA corresponding to the starch binding domain (SBD) of Glucoamylase 1 (1,4-α-D-glucan glucohydrolase) of *Aspergillus niger* (Accession number: X00548) was amplified by PCR from purified genomic

DNA. Primers were designed to allow us to add tails at both ends of the DNA fragment for cloning into pJANGΔCys7Cys236 vector.

Sequence of the PCR primers are:

oligonucleotide 7 (SEQ ID NO: 3):

```
   Gln Ala Cys Thr
5' CAA GCT TGT ACC ACT CCC ACC GCC 3'
    Hind III
```

oligonucleotide 8 (SEQ ID NO: 4):

```
       Ile Cys Arg
5' CCA TAT GCA CCG CCA GGT GTC AGT CAC 3'
    Nde I
```

Amplification, cloning into pGEM-T and sequencing was done as described above. One clone bearing the gene fragment was designated as pGEM-SBD.

Cloning of CM16 and CM17 genes for fusion

Both genes were amplified from purified wheat (*Triticum aestivum*) genomic DNA by PCR.

nucleotide 9 (SEQ ID NO: 5):

5' GTC GGC AAT GAA GAT TGC ACC 3'

nucleotide 10 (SEQ ID NO: 6):

5' TCC AAC TGC GTT CTC CTC TTG GCC 3'

nucleotide 11 (SEQ ID NO: 7):

5' GGA TCC CTA GCT CCA CTG AGA CTC 3'

For CM16 gene (accession number X55455) oligonucleotide 10 and 11, while for CM17 gene (accession number X59791) 9 and 11 pairs were used. Clones were called pGEM-CM16 and pGEM-CM17, respectively. For sub-cloning into pJANGΔCys7Cys236 vector, the genes were PCR amplified again, using pGEM clones as template, purified on QIAquick column and digested with the appropriate enzymes. Primers used in this amplification were:

nucleotide 12 (SEQ ID NO: 8):

```
           Gln Ala Leu Gly
5' TGC GCT CAA GCT TTA GGC AAT GAA GAT TGC ACC 3'
            Hind III
```

nucleotide 13 (SEQ ID NO: 9):

```
                 Ile Cys Ser
5'  CAT ACT CCA TAT GCA GCT CCA CTG AGA CTC 3'
               Nde I
```

Cloning puroindoline A gene for fusion

Lambda genomic clone for puroindoline A (PIN-A), kindly provided by Sadequer Rahman, was used as template to amplify the gene (accession number X69913) by PCR. Primers were designed as follows;

oligonucleotide 14 (SEQ ID NO: 10):

```
   Gln Ala Tyr
5' CAA GCT TAC GAT GTT GCT GGC GGG 3'
    Hind III
```

oligonucleotide 15 (SEQ ID NO: 11):

```
5' CCA TAT GCA CCA GTA ATA GCC AAT AGT GC 3'
    Nde I
```

PCR product was purified, ligated into pGEM-T and sequenced as described above. One clone was used and designated as pGEM-PIN-A.

Fusion of genes or gene fragments with ANG molecule

pJANGΔCys7Cys236 vector was cut with NdeI and HindIII restriction enzymes as all the other pGEM clones and purified on QIAquick columns. Ligation was performed at 14^0C overnight in 10 μl of solution containing T4 DNA ligase and insert:vector at about 20:1 molar ratio. One μl of ligation mix was used for transformation of *E. coli* competent cells and spread onto LB plate with kanamycin. Three fusion gene containing colonies were picked up from each transformation for plasmid preparation and sequencing. The clones were called, for example, pJANG/OHBD/Cys7Cys236 in case of the oleosin binding domain containing ANG molecule.

The fused genes were subcloned into pET-11d expression vector between NcoI and BamHI sites. These clones were called, for example, pET-ANG/OHBD/Cys7Cys236.

Sequence of PCR primers for detection of genes for ANG/domain/Cys7Cys236 in recombinant wheat

The primers used for the detection of genes for ANG/domain/Cys7Cys236 in recombinant wheat are set out below.

Primer	Length	Sequence	Designation
Bx17_3'	23	CAACCATGTCCTGAACCTTCACC	SEQ ID NO:12
RGS2024	18	TGGCTGTTGAGGTTGCAC	SEQ ID NO:13

Expression of fusion proteins

For protein expression one of the pET clones were transformed into *E. coli* strain AD494(DE3) one day before expression work started.

Small scale expression was carried out in 5 ml 2YT medium, supplemented with ampicillin (100 mg/l) using one transformant colony. After about 5 hours from inoculation (OD600=0.4), expression of the protein was induced by addition of 0.4 mM isopropyl-β -D-thiogalactopyranoside (IPTG). Both induced and uninduced cultures were further incubated for 4 hours at 37^0C. Expression was monitored by SDS-PAGE according to Laemmli (1970).

Large scale expression was performed in a shaking flask. One litre 2YT medium was inoculated with 1 ml of overnight culture and induced to express protein by addition of 0.4 mM IPTG at a cell density of ~0.6 Ab. The culture was incubated with shaking overnight and cells were harvested by centrifugation.

Detection of proteins

PAGE gels were stained overnight with 0.025% Coomassie Blue R-250 in 10% TCA. Excess stain was washed away by water-ethanol-acetic acid (8:1:1) solution.

Immunological detection of the PIN-A containing fusion protein was carried out using the method of Ciaffi et al., (1999). The antibody was raised against puroindoline crude extract, kindly provided by Sadequr Rahman.

Other chimeric proteins were detected in immunoblots using an antibody raised against Hordein, kindly provided by John Skerritt.

Purification of the fusion proteins

Expressed ANG/CM16/Cys7Cys236, ANG/CM17/Cys7Cys236 and ANG/SBD/Cys7Cys236 proteins were purified following the method published elsewhere (Tamas et al., 1998), except for the precipitation step. In this work 2 volumes of 1.5 M NaCl were mixed with the 70% ethanol extract, rather than 4 volumes of acetone.

Detection of the gene encoding modified seed storage proteins in transgenic wheat plants

The presence of the gene for modified proteins in transgenic wheat plants was determined by polymerase chain reaction (PCR). Reactions were carried out in 11.6 μl volume containing 9 μl PCR Supermix (GibcoBRL), 50 ng template DNA (extracted from wheat leaf tissue using standard protocols), 172 nmol of each of primers Bx17_3' and RGS2024, and 0.6 μl of 25 mM $MgCl_2$. The PCR conditions were 1 cycle of 94°C for 2 min; 35 cycles of 94°C for 30 sec, 60°C for 30 sec, 72°C for 1 min; 1 cycle of 72°C for 4 min, 25°C for 1 min. The results for one plant containing the gene for ANG/OHBD/Cys7Cys236 are shown in Figure 17. The PCR product of approximately 600 base pairs is indicative of the presence of the gene. A 600 bp PCR product is not obtained in PCR reactions from negative control plants.

RESULTS

Design, construction and cloning of pJANGΔCys7Cys236

The gene selected to construct this vector for fusion protein coded for C hordein. This molecule is a storage protein from barley endosperm and characterised by an absence of cysteine residues. The barley genomic clone encodes a molecule of 261 residues, including a 20 residue signal peptide. The gene for mature protein (molecular mass 28 kDa) has 723 nucleotides, including a 669 bp long fragment for a central repetitive domain. The oligonucleotides for PCR were designed to reduce the size of the central part and substitute one residue by cysteine in both unique terminal domains. Oligonucleotide 1 binds to the 5' end of the gene and has an additional sequence at the 5' end to incorporate an initiation ATG codon for methionine and a restriction site for BspHI. Oligonucleotide 1 has a T at position 22 to replace an A to change the codon of serine to cysteine at position 7. The 3' end of oligonucleotide 2 is complimentary to a sequence in the repetitive domain between 430 and 448 nucleotides. This is a rather unique sequence

within the strongly repetitive region and also codes for an end of a repeat motif. The amplified gene contains only a 411 bp (137 amino acids) long fragment for central repetitive domain. This oligonucleotide also contains the whole sequence of the C-terminal unique region (6 amino acids) of C hordein and a restriction site for BamHI, immediately after the stop codon. To change the threonine residue at the position of 236 of the full size molecule (six residues from the C-terminal end) the oligonucleotide has a C at position 26 and an A at position 27 to substitute a G and a T, respectively. There is another base pair change in this primer at position 22 to replace C with A, which is a "wobble" base of the isoleucine codon. This substitution allowed the creation of restriction site for NdeI enzyme. Oligonucleotide 2 has also got six nucleotides which are not part of the C hordein gene, coding for alanine and serine. These extra nucleotides were added to create one unique restriction site, close to NdeI, within the ANG gene. The two restriction sites are separated with 4 bp to give easy cleavage for both enzymes. The insertion of a gene for another molecule or fragment of a molecule between HindIII and NdeI enabled the present inventors to create fusion proteins with new, designed characteristics.

The gene for ANGΔCys7Cys236 molecule is 474 bp long and codes for a protein with a molecular mass of 18.5 kDa (Figure 3).

PCR amplified DNA was cloned, with two steps, into a plasmid, called pJKKm with a modified MCS. Having deleted HindIII restriction site from the original plasmid, the cloning vehicle pJANGΔCys7Cys236 has unique HindIII and NdeI cleavage sites within ANG gene for gene insertion. Size of the vector is 3873 bp and provides kanamycin resistance to host *E. coli* cells. The gene is sitting in the vector with 5' ends close to the SP6 RNA Polymerase transcription initiation site.

Subcloning engineered genes into pET-11d did not require agarose gel purification of the fragments, because of the difference in resistance genes within plasmids.

Expression of ANGΔCys7Cys236 protein

To check and compare characteristics of this short molecule to hordeinCys7Cys236, it has been expressed in both small and large quantities. Comparison of the SDS-PAGE patterns of the total cell proteins before and 3 hours after induction with IPTG showed a new band in the induced sample. ANG molecule was readily extracted from lysed cells with 70% (v/v) ethanol

and precipitated by the addition of 2 volume of 1.5 M NaCl solution. The resulting preparation, in the presence of reducing agent (0.1 M DTT), had the same mobility as the extra band in *E. coli* lysate. Apparent molecular mass of the protein was about 21 kDa, having a slightly lower mobility in the gel. This characteristic is not unusual for storage proteins. In the absence of DTT, the ethanol extracted sample gave a ladder of bands indicating that the protein is able to form long chains through disulfide bonds.

The ability of the ANG molecule to incorporate into the gluten matrix of the dough was confirmed by a series of mixing experiments carried out using small scale testers.

Results of SDS-PAGE and mixing experiments showed clearly that ANGΔCys7Cys236 protein had the same or similar properties as the 2 cysteine residue containing C hordein protein.

Design, construction and cloning of synthetic gene for oleosin hydrophobic binding domain (OHBD)

Amino acid sequence was designed according to a comparison of four molecules (one maize, two barleys and one rice), using Genetic Computer Group (GCG) program, called "pileup". OHBD gene fragment, for this work, contain the entire region of the lipophilic stretch of oleosin (Figure 2), very similar to barley gene (accession number: X82677). Codon usage was designed to avoid long stretch of Gs and Cs, which could have led to mis-annealing and sequencing problems. The 5' flanking region of the gene fragment consisted of a tripeptide sequence, of the amphipathic N terminus, while the 3' end another tripeptide of the C terminus of oleosin. The central anti-parallel beta stranded domain had 71 residues. The turn of the anti-parallel consists of 13 residues and is the most conserved region. The two anti-parallel strands are highly symmetrical in the pairing of residues of similar hydrophobicity on the opposite strands. These characteristics are very similar to those reported for maize oleosin protein (Huang 1996).

The synthetic gene fragment, flanked with two appropriate restriction sites for subcloning, was cloned into pGEM-T vector. One clone, containing the correct sequence of a fragment of 243 bp nucleotides, was designated as pGEM-OHBD. Size of OHBD fragment was 77 amino acids and had a molecular mass of 7 kDa.

Cloning of starch binding domain (SBD)

Glucoamylase 1 from *Aspergillus niger* comprises two domains, one being a catalytic domain (1-470 residues) and the other (509-616 residues) being responsible for binding granular starch (Le Gal-Coeffet et al., 1995).

DNA corresponding to the SBD was amplified by PCR method from *Aspergillus niger* purified genomic DNA, cloned into pGEM-T vector, sequenced and designated as pGEM-SBD. This clone had a 337 bp long fragment (Figure 4) with two restriction sites for insertion into pJANGΔCys7Cys236 vector and an extra alanine before the very first cysteine of SBD. This residue derived from HindIII restriction site. The binding domain had 108 amino acids, including two cysteines, with a molecular mass of 11.9 kDa.

Cloning of CM16 and CM17 gene for fusion

Both CM16 and CM17 (CM refers to Chloroform/Methanol soluble) proteins, reported as members of the α–amylase/trypsin inhibitor family and also reported that specific lipids are tightly bound to the fraction, were purified from wheat (Kobrehel and Sauvaire, 1990). These two molecules are very similar on the amino acid level, but there are a few differences in the distribution of charged residues (Figure 5).

To clone these two genes, three primers were used to amplify them in two, separate PCR reactions from wheat genomic DNA. One primer (oligonucleotide 11) hybridised to the 3' end of the genes, while to distinguish between the two genes, two specific primers were designed for the 5' ends. pGEM-CM17 clone carried the gene for only the mature protein. However, pGEM-CM16 clone had a few extra base pairs from the signal peptide region. Both clones had a DNA fragment encoding mature chloroform/methanol soluble proteins with 10 cysteine residues and a molecular mass of 13.4 kDa.

To subclone these genes into ANGΔCys7Cys236 carrying vector, two restriction sites were added to one of each ends by PCR, using the same primer pairs for both pGEM clones, as templates. Primer corresponding to the N-terminus of the mature proteins contained nucleotides for HindIII restriction enzyme. It also had an extra alanine residue, and a mutation in the first codon (TTA for leucine), substituting valine (GTC) in CM17 and isoleucine (ATC) in CM16 molecule, because of the sequence requirement for

HindIII enzyme. The fragment cloned into ANG cloning vehicle had 360 nucleotides.

Cloning puroindoline A (PIN-A) gene for fusion

The clone for PIN-A protein, which is capable of binding tightly to both wheat phospholipids and glycolipids (Dubreil et al., 1997), was kindly provided by S. Rahman, derived from a genomic library.

PCR amplified DNA encoded for a 121 amino acid long fragment of puroindoline A, 5 residues shorter in the N terminal region than the mature protein. This shorter protein had 10 cysteine residues and a molecular mass of 14.3 kDa (Figure 6). For insertion into ANG molecule, the pGEM-PIN-A clone also carried two restriction sites (HindIII and NdeI) and an extra alanine, because of HindIII site requirement.

Expression of ANG/OHBD/Cys7Cys236 protein

Having inserted OHBD gene fragment into pJANGΔCys7Cys236 cloning vehicle, the entire gene was subcloned into expression vector pET-11d. Protein was expressed in *E. coli*, using AD494(DE3) strain. This genetically engineered new gene (717 bp long) coded for a protein with a molecular mass of 25.5 kDa.

Cells were harvested by centrifugation after overnight expression (5 ml culture) and resuspended in 100 μl gel loading buffer (0.125 M TRIS/HCl pH 6.8, 4% SDS, 10% (v/v) glycerol), containing 0.1 M DTT. Proteins were analysed by SDS-PAGE and stained with Coomassie blue (Figure 7). Lane A contains standard molecular weight markers, lane B shows a control in which the host *E. coli* cell contains the pET-11d vector, and lane C shows the expression of the recombinant ANG/OHBD/Cys7Cys236 protein (migrating slightly further than the 30 kDa marker) from pET-11d containing the ANG/OHBD/Cys7Cys236 gene. The apparent molecular mass of the newly synthesised protein (~29 kDa) was larger than that calculated on the basis of the nucleotide sequence of DNA. A similar characteristic has been observed previously for different prolamines of cereals. This fusion protein consisted of part of a prolamine and a part of an oleosin molecule. C hordein had been considered as a rod shaped molecule, while oleosin hydrophobic binding domain has anti-parallel strands, which would penetrate into the oil body and thus anchor the protein stably. It means that none of the two fragments has globular structure as compared with the molecular weight marker molecules. Discrepancy in apparent molecular mass on SDS-PAGE was

probably due to these two structural phenomena and it was manifested in lower mobility. Figure 8 shows a Western blot of an SDS-PAGE gel (run in the absence of reducing agents such as dithiothreitol) using antibodies to the C-hordein protein. Lane A contained standard molecular weight markers, lane B contained extract of an *E. coli* cell expressing a modified C-hordein gene containing a single cysteine in the N-terminal region, and lane C contained extract from *E. coli* containing the ANG/OHBD/Cys7Cys236 gene in the pET-11d expression plasmid. The strong antibody cross reaction to hordein antibodies at 29 kDa confirms that the 29 kDa protein is the product of the ANG/OHBD/Cys7Cys236 gene.

Expression and analysis of ANG/PIN-A/Cys7Cys236 protein

A DNA fragment comprised of ANGΔCys7Cys236, and Puroindoline-A genes with a length of 840 nucleotides from kanamycin resistant plasmid (see above) was subcloned into pET-11d expression vector. One transformant colony was picked up and transferred into 5 ml of 2YT medium. Protein expression was induced by IPTG and after 2 and 6 hours cells were harvested by centrifugation and resuspended in 200 μl of gel loading buffer. SDS-PAGE analysis of protein content of bacteria, before and after induction, is shown on Figure 10. Lane A contained standard protein molecular mass markers. Lane B contained extract of cells harbouring the control pET-11d plasmid, lanes C and D contained extracts of cells harbouring the pET-11d vector containing the ANG/PIN-A/Cys7Cys236 gene. Lane C contained extract 2 hours after IPTG induction, lane D contained extract prepared 6 hours following IPTG induction. An extra protein band appeared on gel with an apparent molecular mass of 35 kDa, after induction (lane D). This expressed protein positively reacted with puroindoline A antibody (Figure 11). Lanes A and B contained extract of cells harbouring pET-11d containing the ANG/PIN-A/Cys7Cys236 gene, 2 hours and 6 hours after induction respectively. Lane C contained a Western blot of cells harbouring a plasmid containing the puroindoline A gene alone (not inserted into the ANGΔCys7Cys236 gene), and lane D contained extract of cells harbouring the control pET-11d vector. All lanes were reacted with anti- puroindoline A antibody. Figure 12 shows a Western blot of an SDS-PAGE gel reacted with anti-hordein antibodies. Lane A contained molecular mass markers, lane B contained extract of cells harbouring the control pET-11d vector, lane C contained extract of cells harbouring a plasmid which contains the

puroindoline A gene alone, and lane D contained extract of cells harbouring pET-11d vector containing the ANG/PIN-A/Cys7Cys236 gene. The expressed protein is clearly seen in lane D.

Analogue glutenin-puroindoline A fusion protein had a slightly lower mobility on SDS-PAGE as it was calculated (32.8 kDa) on the basis of amino acid sequence. This small difference was probably due to the rod shape of the ANG part of the molecule. Expressed puroindoline A protein, however, had the same mobility as it was calculated.

Expression of modified proteins following wheat transformation

Microprojectile bombardment is currently the most widely applied technique used to transfer genes into wheat. The transfer of genes into wheat and their expression is carried out using specific DNA constructs containing a selectable marker gene and the gene-of-interest respectively. The selectable marker gene used was contained on the plasmid pEmuKON (Chamberlain et al., 1994). Constructs for the gene-of-interest constructs are shown in Figure 13. Wheat transformation was carried out using paromomycin selection according to the procedures outlined in Witrzens et al., 1998).

Expression, purification and analysis of ANG/SBD/Cys7Cys236 protein

PET11d and the derivative of that plasmid containing the gene for ANG/ SBD/Cys7Cys236 protein (pET-SBD) were separately transformed into strain AD494(DE3). One transformant colony of each was used to inoculate 5 ml cultures of LB media containing 200 mg/l ampicillin. After 4 hours shaking at 37°C these cultures were used to inoculate 600 ml of the same media in 2 l flasks, and these were incubated with shaking overnight (16 hours) at 37°C. IPTG was added to these cultures 6 hours after inoculation to induce expression of protein from the plasmids. Cells from each culture were harvested by centrifugation at 10,000 g for 15 min at 5°C.

One half of the cells from each culture were lysed by incubation for 30 min at room temperature in 35 ml of 50 mM Tris-Cl buffer pH 7.5 containing 8 M urea, 5 mM EDTA and 1 mM DTT. Insoluble material was removed by centrifugation at 2×10^5 g for 1 hr at 15°C. The supernatants were then extensively dialysed against the extraction buffer minus urea (50 mM Tris-Cl pH 7.5, 5 mM EDTA, 1 mM DTT) at 5°C. The cloudy precipitates were removed by centrifugation at 2×10^5 g for 1 hr at 5°C and the supernatants (35 mL) stored at -20°C.

β-Cyclodextrin-binding proteins were purified from 20 ml of each of the two dialysed and clarified bacterial extracts by passage through a 5 ml β-cyclodextrin-sepharose column equilibrated with 50 mM Tris-Cl pH 7.5, 100 mM NaCl, 40 mM dithiothreitol, 1 mM EDTA (column buffer). After loading, the column was washed with 20 ml column buffer and bound proteins eluted with 25 ml column buffer containing 15 mM β-cyclodextrin. Five ml fractions were collected during loading, washing and elution of the column. Fractions were dialysed extensively against water, lyophilised and resuspended in 0.1 M Tris-Cl pH 6.8 at ca. 1 μl per ml equivalent of the original bacterial cell culture.

Native- and SDS-PAGE gels of the final load fraction and the wash and elution fractions are shown in Figure 14. Each lane contained protein from the equivalent of 7 ml of the bacterial culture. Two major β-cyclodextrin-binding proteins were seen in the elution fractions (see lanes 6 and 7) from the strain bearing plasmid pET-SBD. These are estimated to be 30 kDa and >60 kDa from the SDS-PAGE gel. A 30 kDa protein was also eluted from the column of the extract from the control strain bearing plasmid pET11d (panel C, lane 6), but it is not produced to the same level as in the former strain.

Expression, purification and analysis of ANG/CM16/Cys7Cys236 protein

The gene for this fusion protein (834 bp long) was subcloned, from kanamycin resistant, pJKKm originated plasmid, into the ampicillin resistant pET-11d expression vector, as above. Bacteria was grown at 37^0C until OD600 reached 0.6 units, then expression was induced by adding IPTG. Cells were harvested two, four or six hours after induction and samples were prepared as in case of SBD chimeric protein.

Result of the SDS-PAGE analysis is shown in Figure 9 (the gel is stained with Coomassie protein stain). Lane A contains standard protein molecular weight markers. Lanes B to E show ethanol-soluble extracts of the crude *E. coli* lysates. Lane B contained extract from cells containing the control plasmid, pET-11d. Lanes C, D and E contained ethanol soluble extracts of cells harbouring the pET-11d vector containing the ANG/CM16/Cys7Cys236 gene, prepared from cells 2, 4 and 6 hours respectively after induction of protein synthesis using IPTG. The ethanol-soluble fraction gave only one band on Coomassie stained gel representing a protein with an apparent molecular mass of 35 kDa. This was also higher

than the calculated (31.9 kDa), and this discrepancy is also explained by the unusual shape of the protein.

Yield of protein expression in large scale, in shaking flask was 30 mg protein from 1 litre of medium.

Demonstration of Starch Binding Activity

The starch-binding activities of the major proteins in fractions 6 and 7 of the affinity purification column were demonstrated as follows. Fractions were loaded onto four native-PAGE gels containing glycerol (10% w/v), CHAPS (0.05% w/v) and various combinations of β-cyclodextrin (5 mM) and starch (corn amylopectin, 0.1% w/v), as indicated in Figure 18. The mobility of the protein bands in each gel was compared. It was found that the electrophoretic mobility of the major protein band in each fraction was significantly retarded in the presence of starch (Figure18, panel B). The specificity of the starch binding activity of these protein bands was demonstrated by its competitive inhibition by β-cyclodextrin (panels C, D).

Identification of the purified starch-binding proteins by tryptic-peptide mass fingerprinting and N-terminal protein sequencing

The two major protein bands from the native gel of β-cyclodextrin affinity column fractions from strain AD494(DE3)/pET-SBD were subjected to in-gel tryptic digestion followed by matrix-assisted laser-desorption time-of-flight (MALDI-TOF) mass spectroscopy. The masses of the identified tryptic fragments were compared to the theoretical masses of tryptic peptides expected from ANG/SBD/Cys7Cys236, as well as being used to search the protein sequence databases for other candidate proteins giving matching tryptic peptides.

The 30 kDa protein of fraction 6 (Figure 14, panels A, B, lane 6) was identified as the periplasmic maltose-binding protein of *E. coli* (product of the *malE* gene), based on 17 peptide matches covering 57% of the protein sequence.

The major protein band of fraction 7 (Figure 14, panels A, B, lane 7) was identified as ANG/SBD/Cys7Cys236 on the basis of 4 matching peptides covering 23.11% of the sequence (Table 1). Further proof of the identity of this protein was obtained from N-terminal protein sequencing. Ten cycles of sequencing yielded the sequence:

M R Q L N P C S Q E

which is as expected for ANG/SBD/Cys7Cys236. This provides convincing proof of the identity of the purified protein, and confirms that the functional data are properties of the modified seed storage protein, as designed.

Table 1. Mass spectral characterisation of ANG/SBD/Cys7Cys236

Peptide mass	Expected mass	Start	End	Sequence
1776.74	1777.28	224	238	IESDDSVEWESDPNR
2193.00	2193.85	221	238	FIRIESDDSVEWESDPNR
2652.26	2653.54	198	220	YTSSDPLWYVTVTLPAGESFEYK
2257.03	2258.00	239	258	EYTVPQAC*GTSTATVTDTWR

C* = acrylamide adduct of cysteine

Demonstration of the ability of ANG/SBD/Cys7Cys236 to form disulphide-bonded polymers

The design of the analogue glutenin protein (ANG) included cysteine residues at both N- and C-termini to allow disulphide-mediated incorporation of the modified seed-storage protein into the gluten macropolymer. The present inventors have shown by *in vitro* polymerisation experiments that the ability of the ANG/domain/Cys7Cys236 proteins to incorporate into disulphide-bonded polymeric species is not impaired by the presence of the extra domain.

The ability to form polymers was demonstrated by copper-mediated oxidation of ANG/SBD/Cys7Cys236 protein in 50 mM Tris-Cl buffer pH 6.8. Three pairs of protein samples were employed, with one of each pair containing added β-cyclodextrin (0.5 mM) to analyse any effect of a ligand on the disulphide-bonding of the ANG. Two pairs were oxidised for 5 min at room temperature in the presence of 0.1 mM $CuSO_4$, before chelation of the metal by addition of EDTA (10 mM). One of these pairs was then reduced in the presence of 20 mM DTT. A further pair of samples had no copper added, but were reduced with DTT in the presence of EDTA. The samples were analysed by SDS-PAGE (Figure 19). There was no effect on oxidation due to the presence of the ligand β-cyclodextrin. The majority of the oxidised protein sample forms a high molecular weight band which just enters the separating gel (lanes 1, 2). That this change in mobility is due to the

formation of disulphide-bond polymeric forms is shown by the ability of the disulphide-specific reductant DTT (lanes 3, 4) to break the polymer back down into the higher-mobility form seen in the reduced sample (lanes 5, 6).

Detection of the gene encoding modified seed storage proteins in transgenic wheat plants.

The presence of the gene for modified proteins in transgenic wheat plants was determined by polymerase chain reaction (PCR). Reactions were carried out in 11.6 μl volume containing 9 μl PCR Supermix (GibcoBRL), 50 ng template DNA (extracted from wheat leaf tissue using standard protocols), 172 nmol of each of primers Bx17_3' and RGS2024, and 0.6 μl of 25 mM $MgCl_2$. The PCR conditions were 1 cycle of 94°C for 2 min; 35 cycles of 94°C for 30 sec, 60°C for 30 sec, 72°C for 1 min; 1 cycle of 72°C for 4 min, 25°C for 1 min. The results for one plant containing the gene for ANG/OHBD/Cys7Cys236 are shown in Figure 17. The PCR product of approximately 600 base pairs is indicative of the presence of the gene. A 600 bp PCR product is not obtained in PCR reactions from negative control plants.

Demonstration of the ability of ANG/SBD/Cys7Cys236 to form disulphide-bonded polymers

The design of the analogue glutenin protein (ANG) included cysteine residues at both N- and C-termini to allow disulphide-mediated incorporation of the modified seed-storage protein into the gluten macropolymer. The present inventors have shown by *in vitro* polymerisation experiments that the ability of the ANG/domain/Cys7Cys236 proteins to incorporate into disulphide-bonded polymeric species is not impaired by the presence of the extra domain.

The ability to form polymers was demonstrated by copper-mediated oxidation of the protein in 50 mM Tris-Cl buffer pH 6.8. Three pairs of protein samples were employed, with one of each pair containing added β-cyclodextrin (0.5 mM) to analyse any effect of ligand-binding on the disulphide-bonding of the ANG moiety. Two pairs were oxidised for 5 min at room temperature in the presence of 0.1 mM CuSO4, before chelation of the metal by addition of EDTA (10 mM). One of these pairs was then reduced in the presence of 20 mM DTT. A further pair of samples had no copper added, but were reduced with DTT in the presence of EDTA. The samples were analysed by SDS-PAGE (Figure19). There was no effect on oxidation due to

the presence of the ligand β-cyclodextrin. The majority of the oxidised protein sample forms a high molecular weight band which just enters the separating gel (lanes 1, 2). That this change in mobility is due to the formation of disulphide-bond polymeric forms is shown by the ability of the disulphide-specific reductant DTT (lanes 3, 4) to break the polymer back down into the higher-mobility form seen in the reduced samples (lanes 5, 6).

SUMMARY

The demonstration of the binding activity of a modified glutenin according to the present invention was provided in the example given above for what has been termed the ANG/SBD/Cys7Cys236 protein. The data provided supports the demonstration of the expression and purification of a novel protein (apparent molecular mass 60 kDa) from the *E. coli* strain bearing the expression plasmid (pET-SBD), in comparison to the lack of this protein in the same strain bearing the control plasmid (pET11d) (Figure 14). The purification involved the use of an affinity chromatography step in which the protein was purified on the basis of its ability to bind to a column matrix containing β-cyclodextrin. The protein was specifically eluted from this column using unbound β-cyclodextrin. This, therefore, represents the first example of how the present inventors demonstrate reduction to practice through the generation of a seed storage protein containing a functional ligand binding domain.

The demonstration of the binding of the major purified proteins to corn starch using native polyacrylamide gel electrophoresis, and the inhibition of this binding by β-cyclodextrin (PAGE) (Figure18) is a highly convincing demonstration of reduction to practice of the present invention.

The identification of this protein as ANG/SBD/Cys7Cys236 by peptide mass fingerprinting and N-terminal protein sequencing further demonstrates the functionality of modified seed-storage proteins containing a macromolecular binding domain. The anomolous apparent molecular weight of this protein in SDS-PAGE may be yet another example of the general phenomenon of modified C hordein proteins migrating with unexpected electrophoretic mobility, as the calculated mass of the protein is only 30 kDa. Alternatively, the protein may form a dimer, perhaps around a single bound b-cyclodextrin ligand, which is not dissociated by boiling in the presence of SDS and a reducing agent.

The ANG portion forms disulphide-bonded polymeric species was demonstrated by native-PAGE of the copper-oxidised affinity purified proteins. The fusion protein has been demonstrated to contain the novel starch binding activity. Importantly, it is believed that the modified protein has not lost the ability to polymerise through the cysteine residues at the N- and C-termini and therefore the novel protein can be incorporated into the gluten macropolymer in wheat flour.

Demonstration of the capacity of the modified seed-storage protein with macromolecule-binding domain (in this instance ANG/SBD/Cys7Cys236) to form disulphide-bonded polymeric species (Figure 19). This demonstrates the ability of the protein to be incorporated into the gluten macropolymer through disulphide bonds.

Table 1 demonstrates the identity between the molecular weights of the experimentally determined tryptic fragments of the major β-cyclodextrin-binding protein (starch-binding protein) purified from the ANG/SBD/Cys7Cys236 strain and the theoretical masses of the tryptic fragments of the ANG/SBD/Cys7Cys236 protein. The identification of the 30kDa protein as an endogenous bacterial maltose-binding protein is also provided.

It will be appreciated by persons skilled in the art that numerous variations and/or modifications may be made to the invention as shown in the specific embodiments without departing from the spirit or scope of the invention as broadly described. The present embodiments are, therefore, to be considered in all respects as illustrative and not restrictive.

REFERENCES

Aalen, R.B. 1995. The transcript encoding two oleosin isoforms are both present in the aleurone and in the embryo of barley (Hordeum vulgare) seeds. Plant Mol. Biol. 28:583-588

Bekes, F., Gras, P.W. and Gupta, R.B. 1994. Mixing properties as a measure of reversible reduction and oxidation of doughs. Cereal Chemistry 71:44-50

Bushuk, W. 1998 . Interactions in wheat doughs. *In* Interactions: The keys to cereal quality. R.J. Hamer, R.C. Hoseney Eds. American Association of Cereal Chemists, Inc. St Paul, Minnesota. pp1-16.

Chamberlain, D.A., Brettell, R.I.S., Last, D.I., Witrzens, B., McElroy, D., Dolferus, R., Dennis, E.S. 1994. The use of the EMu promoter with antibiotic and herbicide resistance genes for the selection of transgenic wheat callus and rice plants. Australian Journal of Plant Physiology. 21: 95-112.

Ciaffi, M., Lee, Y.K., Tamas, L., Gupta, R., Skerritt, J. and Appels, R. 1999. The low-molecular-weight glutenin subunit proteins of primitive wheats. III. The genes from D-genome species. Theoretical and Applied Genetics 98: 135-148.

Dubreil, L., Compoint, J-P. and Marion, D. 1997. Interaction of Puroindolines with wheat flour polar lipids determines their foaming properties. J. Agric. Food Chem. 45:108-116

Gan, Z., Ellis, P.R. and Schofield, J.D. 1995. Gas cell Stabilisation and gas retention in wheat bread dough. Journal of Cereal Science 21:215-230

Huang, A.H.C. 1996. Oleosins and oil bodies in seeds and other organs. Plant Physiology 110:1055-1061

Kasarda, D.D. 1989. Glutenin structure in relation to wheat quality. Pages 277-302 in : Wheat is unique. Y. Pomeranz, Ed. Amer. assoc. Cereal Chem., St Paul, MN.

Kirschman, J.A. and Cramer, J.H. 1988. Two new tools: multi-purpose cloning vectors that carry kanamycin or spectinomycin/streptomycin resistance markers. Gene 68: 163-165.

Kobrehel, K. and Sauvaire, Y. 1990. Particular lipid composition in isolated proteins of durum wheat. J. Agric. Food Chem. 38:1164-1171

Laemmli, U.K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227:680-685

PCT/AU99/00563
Received 1 June 2000

Le Gal-Coeffet. M-F., Jacks. A.J.. Sorimachi.K.. Williamson. M.P., Williamson. G. and Archer. D.B. 1995. Expression in Aspergilus Niger of the starch-binding domain of glucoamylase. Eur. J. Biochem. 233:561-567

MacRitchie. F. 1992. Physicochemical properties of wheat proteins in relation to functionality. Adv. Food Nutr. Research 36:1-87

Morrison, W.R. 1989. Recent progress on the chemistry and functionality of flour lipids. Pages 131-149 in: Wheat end-use properties: Wheat and flour characterisation for specific end-uses. H. Salovaara, Ed. University of Helsinki, Lahti.

Tamas, L., Bekes. F.. Greenfield, J., Tatham, A.S.. Gras, P.W.. Shewry, P.R. and Appels, R. 1998. Heterologous expression and dough mixing studies of wild-type and mutant C hordeins. Journal of Cer. Science 27:15-22

Weeks, J.T.. Anderson, O.D. and Blechl, A.E. 1993. Rapid production of multiple independent lines of fertile transgenic wheat (Triticum aestivum). Plant Physiology 102:1077-1084

Witrzens. B.. Brettell. R.I.S., Murray, F.R.. McElroy, D., Li. Z., Dennis. E.S. 1998. Comparison of three selectable marker genes for transformation of wheat by microprojectile bombardment. Australian Journal of Plant Physiology 25:39-44.